Marine Resource Economics, Volume 22, pp. 155–172
0738-1360/00 $3.00 + .00

Accommodating Complex Substitution Patterns in a Random Utility Model of Recreational Fishing

LEN M. HUNT
Ontario Ministry of Natural Resources

PETER C. BOXALL
University of Alberta

BARRY BOOTS
Wilfrid Laurier University

Abstract *We employed a cross-nested logit (CNL) model that permits a rich pattern of substitution among alternatives within a closed form choice model. The specification we employed is ideal for applications with many choice alternatives, such as the 431 fishing sites in this application. The CNL model provided a significant improvement over multinomial and nested logit model specifications at explaining observed recreational fishing site choices by residents of northern Ontario, Canada. Results from two scenarios illustrated the implications of using the CNL model on spatial substitution patterns and welfare measures associated with attribute change scenarios. The CNL model forecasts demonstrated that the relative change in fishing site use was lower at the most affected sites and higher at sites near the affected sites than was predicted by the multinomial logit model. No consistent pattern was found in mean or variance of welfare estimates associated with hypothetical attribute changes.*

Key words Compensating variation, cross-nested logit, fishing site choice, random utility model, spatial substitution.

JEL Classification Code Q26.

Introduction

Random utility models permit researchers to estimate changes in economic welfare associated with different fisheries management scenarios. These models focus on individual angler site choices by developing indirect utility functions for alternative

---

Len M. Hunt is a research scientist at the Centre for Northern Forest Ecosystem Research, Ontario Ministry of Natural Resources, 955 Oliver Rd. Thunder Bay, ON Canada, P7B 5E1, email: len.hunt@mnr.gov.on.ca. Peter C. Boxall is a professor in the Department of Rural Economy, University of Alberta, 114 St. – 89 Ave., Edmonton, AB Canada, T6G 2E1, email: pboxall@ualberta.ca. Barry Boots is a professor in the Department of Geography and Environmental Studies, Wilfrid Laurier University, 75 University Avenue, West, Waterloo, ON Canada, N2L 3C5, email: bboots@wlu.ca.

We thank two anonymous reviewers for their constructive and insightful suggestions on ways to improve this paper. We also thank Sarah Browne, Jeff Moore, and Mandie Ross for overseeing field data collection. The Living Legacy Trust, Ontario Ministry of Natural Resources, Ontario Federation of Anglers and Hunters, and the Northwestern Ontario Sportsmen's Alliance provided funding for this research. Much of the work on this manuscript was completed while Boxall was a Research Fellow at the Policy and Economics Research Unit at CSIRO Land and Water, Adelaide, Australia, and he is grateful for their support.

fishing sites that are conditional on their choice by the angler. These functions consist of observed or deterministic components related to site and angler attributes and an unobserved or stochastic component. Besides welfare measures, an important contribution of these models is an assessment of substitution effects among alternatives. The degree of substitutability among alternatives can affect forecasts of choice and economic welfare from scenarios. However, most recreation-based economic applications have not examined methods that account for complex substitution patterns among choice alternatives (Herriges and Phaneuf 2002). Many empirical applications employ statistical models, such as the multinomial logit, that impose rigid patterns of substitution among alternatives through the independence of irrelevant alternatives (IIA) property.

The IIA property is a consequence of assuming that the differences in unobserved utilities among alternatives are independently and identically distributed. Consequently, the attributes and parameter estimates within the deterministic utility for alternatives capture all necessary substitution effects among the alternatives.

While the IIA property may apply to specific choice applications, spatial proximity of alternatives may lead to significant shared unobserved utility among the alternatives. From a nuisance perspective, spatial relationships among the unobserved utilities may arise from model misspecification and taste variation. For example, the omission of an important site attribute (*e.g.*, congestion) that has a non-random spatial pattern of values will lead to some sharing of unobserved utilities among spatially near sites. Spatial relationships may also arise from taste variations when researchers restrict individual preferences for site attributes to be equal. If these preferences vary and the site attributes have a non-random spatial pattern of values, some sharing of unobserved utilities among spatially near sites will occur. Space may also lead to sharing among unobserved utilities from a substantive process. Recreationists may become attached to places or may narrow choices by first selecting a region followed by a site within the chosen region. From either perspective, the assumption of independence of unobserved utilities among the alternatives is violated.

Advanced generalized extreme value (GEV) (McFadden 1978) and open form choice models, such as mixed logit, provide two ways that researchers can account for complex patterns of substitution among unobserved utilities. A nested logit is a GEV model that many researchers have employed to relax the IIA property (*e.g.*, Kling and Thomson 1996; Jones and Lupi 1999; Hauber and Parsons 2000; Grijalva *et al.* 2002). A nested logit model separates the unobserved utility of an alternative into unique and shared components. Researchers typically assign the alternatives to one of several nests. Any shared unobserved utility among alternatives within a nest will induce positive correlation among the unobserved utilities. This correlation will result in alternatives within nests acting as better substitutes for each other than alternatives between nests.

While the nested logit model permits different patterns of substitution among alternatives, the IIA property holds for all alternatives within the same nest. The nested logit model also requires researchers to *a priori* specify alternatives among discrete nests. While discrete nests may make sense for some criteria (*e.g.*, saltwater and freshwater sites), spatial relationships among the unobserved utilities are unlikely to fit into non-overlapping nests.

Other GEV models, such as the generalized nested logit (GNL) (Wen and Koppelman 2001), permit flexibility in estimating relationships among the unobserved utilities of alternatives. The GNL model differs from the nested logit by permitting researchers to allocate alternatives into multiple nests. By allocating alternatives into multiple nests (*i.e.*, cross-nesting), researchers do not impose the IIA property on any alternative. The use of GNL models is growing in transportation research and other areas where there are few choice alternatives (*e.g.*, Bierlaire, Axhausen, and Abay 2001; Wen and Koppelman 2001).

Other researchers have used open form choice models, such as the multinomial probit and mixed logit, to account for spatial substitution effects. By assuming a process for relationships among unobserved utilities (*e.g.*, spatial autoregressive), one can account for relationships among unobserved utilities while limiting the number of integrals required for estimation (*e.g.*, Bolduc 1992, 1999; Bolduc, Fortin, and Fournier 1996; Bolduc, Fortin, and Gordon 1997; and Garrido and Mahmassani 2000). Bolduc, Fortin, and Gordon (1997) also demonstrated that one may estimate all identifiable elements in the covariance matrix of unobserved utilities with Bayesian estimation approaches. Finally, Herriges and Phaneuf (2002) showed the benefits of using a mixed logit model to account for correlations among unobserved utilities for both alternatives and choice occasions.

Many of the models mentioned above are difficult to estimate with large numbers of alternatives (*i.e.*, greater than 50 alternatives). As an exception, Bhat and Guo (2004) used a cross-nested GEV model for a choice problem with many alternatives. These authors examined the housing site choices of individuals from Dallas County, Texas, for 98 aggregated alternatives. A tractable model was created by *a priori* allocating alternatives to nests by spatial proximity (*i.e.*, whether the alternatives were neighbours). The authors also transformed the allocations to ensure that they summed to one for any alternative. The estimated model revealed that alternatives in close spatial proximity shared some unobserved utility. Consequently, the proximity of the alternatives affected the pattern of substitution among the various alternatives.

We employ a similar closed form GEV model to allow for a rich pattern of spatial substitution among alternatives in a recreational fishing application. A cross-nested logit (CNL) model is used to examine the fishing site choices by a resident angling population from northern Ontario, Canada. The CNL model is particularly appealing in this empirical application since there are many angling opportunities due to an abundance of lakes and rivers. This application followed the approach of Bhat and Guo (2004) by *a priori* specifying the number of nests and the allocation of alternatives to the nests based on spatial considerations.

This paper provides three assessments among the CNL, nested (NL), and multinomial logit (MNL) models. First, we examine whether the CNL model provides a statistical improvement over MNL and NL models at fitting observed site choices. Second, we forecast changes in site choices using the CNL and MNL models for two hypothetical scenarios and compare the resulting spatial distribution of changes. Finally, we assess the welfare implications of these forecasts by estimating per-trip compensating variation (CV) for the two scenarios.

Model Formulation

We assume that an angler is a utility maximizer who, when confronted with a set C containing J fishing sites, chooses the site that maximizes his or her utility. Difficulties in understanding and accounting for all aspects that lead to an individual's utility (*e.g.*, preference heterogeneity) create uncertainty when estimating the utility for an individual. Equation (1), which is adapted from Ben-Akiva and Bierlaire (1999), represents a general form of an individual's (n) utility (U) for an alternative (i):

$$U_{in} = \ln\left(\sum_{m=1}^{M} \alpha_{im} e^{V_{in} + \varepsilon_{in} + V_{mn} + \varepsilon_{mn}}\right). \tag{1}$$

Utility depends upon M different nests (groups) of alternatives that researchers identify *a priori*. The terms after the exponent are the systematic (V) and the unobserved (ε) utilities at both the alternative (i) and nest (m) levels. The α term represents the likelihood that alternative i belongs to nest m. Consequently, α weights the contribution of the systematic and unobserved utilities for each nest. The equation demonstrates that the utility of each alternative consists of M contributions.

We simplify equation (1) by using V_{imn} to capture the systematic utility for alternative i and nest m. The probability of selecting alternative i from a set of C alternatives is:

$$P_{in} = \Pr\left\{ \left[\ln\left(\sum_{m=1}^{M} \alpha_{im} e^{V_{imn} + \varepsilon_{in} + \varepsilon_{mn}} \right) \right] \geq \left[\ln\left(\sum_{m=1}^{M} \alpha_{jm} e^{V_{jmn} + \varepsilon_{jn} + \varepsilon_{mn}} \right) \right], \forall j \in C \right\}. \qquad (2)$$

Different assumptions made by researchers about the nests, allocation of alternatives to nests, and the unobserved utilities lead to different statistical models. The multinomial or conditional logit model (see equation [3]) arises when researchers assume there is one nest (M=1), and the difference in unobserved utilities ($\varepsilon_{in} - \varepsilon_{jn}$) are independently and identically distributed as type I extreme value. Since there is only one nest, the unobserved (and systematic) utilities specific to the nest cancel from equation (2). As well, the allocation of any alternative to this one nest reduces equation (1) to $U_{in} = V_{in} + \varepsilon_{in}$. The μ term in equation (3) is inversely related to the variance of the difference in unobserved utilities:

$$P_{in} = \frac{e^{\mu V_{in}}}{\sum_{j \in C} e^{\mu V_{jn}}}. \qquad (3)$$

Another assumption about equation (2) involves some sharing of unobserved utilities (ε_{mn}) among the M different nests. The researcher also assumes that each alternative is allocated to one nest and that he/she knows this allocation with certainty (*i.e.*, the allocation parameters (α) take on values of zero or one). Finally, researchers must make assumptions about the cumulative density function for the unobserved utility. Since many researchers provide detailed expositions of nested and other GEV models, these details are not provided here (Morey 1999; Ben-Akiva and Bierlaire 1999; Train 2003). Instead, we appeal to GEV theory (McFadden 1978) and assume that the generating function for this model is:

$$G(Y_1, Y_2, \ldots, Y_J) = \sum_{m=1}^{M} \left[\sum_{i=1}^{J} (\alpha_{im} Y_{in})^{\frac{1}{\theta}} \right]^{\theta}. \qquad (4)$$

The generating function for the J alternatives depends upon values (Y) for each alternative, allocation (α) of an alternative to each of the M nests, and a dissimilarity parameter (θ).[1] By specifying $Y = e^{V_{imn}}$, and restricting $\alpha > 0$ and $0 < \theta < 1$, equation (4) satisfies the conditions for GEV family membership and is thus consis-

[1] The dissimilarity parameter (θ) arises from normalizing the scale parameter that arises from the ε_m unobserved utility and taking the inverse of the μ parameter, which more generally can take on separate values for each nest (μ_m).

tent with random utility theory (see Wen and Koppelman (2001) for details).

Equation (4) leads to the NL model illustrated in equation (5). The reason for this general illustration of the NL will become apparent. The allocation parameters (α) are the researcher-specified probabilities of alternative *i* belonging to the nest *m*. The dissimilarity parameter (θ) measures the independence of unobserved utility among the alternatives (Train 2003). Equation (5) collapses to the multinomial logit model when $\theta = 1$:

$$P_{in} = \sum_{m=1}^{M} \left\{ \frac{(\alpha_{im} e^{V_{imn}})^{\frac{1}{\theta}} \left[\sum_{j=1}^{J} (\alpha_{jm} e^{V_{imn}})^{\frac{1}{\theta}} \right]^{\theta-1}}{\sum_{l=1}^{M} \left[\sum_{j=1}^{J} (\alpha_{jl} e^{V_{jln}})^{\frac{1}{\theta}} \right]^{\theta}} \right\}. \tag{5}$$

A different model from equations (1) and (2) arises from an assumption of sharing of unobserved utilities among the *M* nests. The researcher does not discretely allocate alternatives to the nests for one of two reasons. First, one may assume that the sharing of unobserved utility among the nests arises from criteria that are not best defined by discrete groups. For example, the effects of space are often better captured by proximity than discrete regions. Second, one may assume that while individuals allocate alternatives discretely into nests, constraints and taste variations may lead different individuals to assign an alternative to different nests. Without observing the reasons for differences in allocation, researchers may estimate models for the average respondent, which requires allocating portions of alternatives to nests. In either instance, the α_{im} term, while deterministic, will take on values that range from zero to one. This CNL model (Bhat and Guo 2004) arises from the generating function of equation (4) and takes the form of equation (5).

Two further generalizations to equations (4) and (5) involve the α and θ parameters. One could estimate the α parameters within the choice model leading to a general model that would capture the NL and the CNL of Bhat and Guo (2004) as special cases. One could also permit different θ parameters for each nest, which would allow a richer pattern of correlation among the unobserved utilities for the alternatives. These further generalizations are described under the generalized NL model of Wen and Koppelman (2001).

The large number of alternatives in this application (greater than 400) limits our focus to the MNL, NL, and CNL models. The systematic utility for each site (V_{imn}) consists only of site attributes and changes in income arising from accessing the sites through travel costs. These attribute measures associated with site *i*, $\mathbf{X}_i$, are integrated with estimated parameters (*i.e.*, preferences) for these attribute measures (β). Most applications, including here, combine the attributes and preferences in a linear additive manner (*i.e.*, $V_{inm} = \mathbf{X_{in}}\beta$).

The *CV* associated with changes in attributes at one or more sites is described by a slightly more general formula than that provided by Morey (1999) for the NL. This *CV* formula with zero income effects is shown in equation (6). The zero and one subscripts denote the deterministic utility levels before and after changes to the sites:

$$CV_n = \frac{1}{\beta_{\cos t}} \left[\ln\left\{ \sum_{m=1}^{M} \left[\sum_{i=1}^{J} (\alpha_{im} e^{\mathbf{X_{0in}}\beta})^{\frac{1}{\theta}} \right]^{\theta} \right\} - \ln\left\{ \sum_{m=1}^{M} \left[\sum_{i=1}^{J} (\alpha_{im} e^{\mathbf{X_{1in}}\beta})^{\frac{1}{\theta}} \right]^{\theta} \right\} \right] \tag{6}$$

When the dissimilarity parameter equals one, equation (6) reduces to the typical CV estimate for the MNL model discussed by Hanemann (1982). The next section describes the development of the data used to estimate the deterministic utility parameters including the travel cost parameter β_{cost}.

Data

The empirical data come from open water fishing sites chosen by anglers residing in Thunder Bay, Ontario, Canada. Research assistants visited and verified the existence of 629 access points that were accessible by roads, trails, or popular portages in the Thunder Bay area. Since some of these access points were located on the same waters, we used 431 fishing sites as alternatives for the fishing site choice models. When reducing the number of fishing sites, the access point with the minimum travel cost was always chosen.

A diary was used to collect information from anglers about chosen fishing sites during the April 1 to September 30 fishing season in 2004. A consultant administered a very short telephone survey to 933 anglers, of which 655 agreed to participate in the angling diary program. The diaries were sent by mail in three two-month waves, and anglers were asked to record fishing trip details from April 1 to September 30, 2004. A total of 347 anglers (53.1%) provided complete trip information, while an additional 53 (8.1%) provided partial trip information over the period.

The 347 anglers from the Thunder Bay area who completed the diary reported taking 2,262 fishing trips over 4,625 days during the April 1 to September 30, 2004 season. We only modeled fishing trips from May 1 to September 30, 2004 to avoid problems of having ice and open water season trips together. Only two types of fishing trip contexts were modelled. First, day fishing trips that were not taken to private accommodation and not part of a longer trip from home were included (representing 39.9% of fishing trips). Second, similar multiple-day trips with added caveats of being less than seven days and for the expressed primary purpose of fishing were also included (representing 11.0% of fishing trips). The final choice models were estimated using information from 1,152 fishing trips.

An obvious problem of having 431 alternatives was to identify a tractable choice model. A series of eight spatial support points, shown in figure 1, were used to limit the number of nests while accounting for sharing of unobserved utilities among the alternatives. These support points were chosen to represent fishing areas as described in personal interviews by local anglers and experts. For the NL model, alternatives were assigned to the nearest spatial support point as measured by road distance. For the CNL model, the allocation of an alternative to a nest was based on the inverse road distance (d) separating the spatial support point (m) and fishing alternative (i) divided by the summation of this measure for all M nests (see equation [7]):

$$\alpha_{im} = \frac{\frac{1}{d_{im}}}{\sum_{l=1}^{M} \frac{1}{d_{il}}}. \tag{7}$$

This deterministic approach yielded considerable cross nesting of alternatives among the nests in contrast to the discrete allocation for the nested logit. While the results are conditional upon the number and orientation of spatial support points in

Figure 1. Spatial Support Points for the Effective* Thunder Bay Study Area
* Sites outside the effective study area were visited but not included in the model estimation.
Source: Hunt (2006).

figure 1 and the form this relationship takes as shown in equation (7), this structure serves to illustrate an approach to estimate a model with hundreds of alternatives. The alternatives of fishing outside the study area and fishing at unknown locations were not allocated to any of the nests based on the spatial support points. The deterministic utility for these alternatives was captured solely through alternative specific constants.

We employed a standard approach to estimate travel costs (*e.g.*, Englin, Boxall, and Watson 1998). This approach involved two calculations. First, out-of-pocket expenses were estimated as vehicle operating costs based on the average operating expense for a Dodge Caravan in 2004 driven 18,000 km annually with fuel costs of $0.744/litre (Canadian Automobile Association 2004). The resulting $1.09 estimate per return kilometre was halved to account for ride sharing among the anglers who typically fish and travel in groups. Travel distances were determined for each angler origin and fishing site through automated network GIS analyses based on minimizing travel time.

Second, the opportunity cost of time was estimated as one-third of the average annual personal income divided by 2,080 (*i.e.*, 52 weeks and 40 hours per week). Since no information about income was collected from the angling diary participants, the average personal income for Thunder Bay residents was used. Typically, a constant speed of travel is used to assist in the calculation of the opportunity cost of time. In this application, however, prominent features associated with recreational fishing are the availability and quality of road access to fishing sites. Since roads and trails in this region are heavily influenced by industrial activity (particularly forest harvesting), the quality of road surfaces can be variable. To capture this effect, travel speeds were adjusted to account for the heterogeneity in road quality. Table 1 provides information about the various types of roads and trails and assumed travel speeds that were used to estimate travel times. Research assistants conducted field inspections of all roads and trails to evaluate road and trail type and estimate travel speeds.

Results

The site choice models employed attributes to capture cost, fishing quality, facility quality, regulations, and site development. The previous section described the travel cost attribute. Table 2 provides the labels and descriptions of these attributes used to estimate the site choice models.

Table 1
Different Road and Trail Classifications and Travel Speeds used to Estimate Travel Times

Road/Trail Type	Lanes	Maintenance	Travel Speed (km/hr)
Paved highway	Two or more	Good	90
Gravel highway	Two	Good	70
Gravel road 1	Two	Good	60
Gravel road 2	Usually two	Adequate to poor	50
Gravel lane 1	One	Adequate	45
Gravel lane 2	One	Poor	30
Gravel lane 3	One	Very poor	15
Trail 1	Trail	None	7.5
Trail 2	Trail	None	5

Table 2
Description of Attributes Included in the Fishing Site Choice Models

Label	Description
θ	Dissimilarity parameter estimate for all nests
OUTSIDE	Alternative specific constant for sites outside the region (1, 0)
UNKOWN	Alternative specific constant for unknown sites within the region (1, 0)
A_WALL	Availability of walleye (0, 1)
A_BASS	Availability of smallmouth bass (0, 1)
A_LTROUT	Availability of lake trout (0, 1)
A_BTROUT	Availability of brook trout (0, 1)
A_BSTR	Availability of smallmouth bass and any type of trout species (0,1)
E(W_CUE)	Estimated walleye catch rate per hour of fishing
RT_CUE	Average reported rainbow trout catch rate per one hour of fishing
LN_WAREA	Natural logarithm of area of fishing waters (ha)
T_COST	Travel costs (vehicle operation and travel time)
PORTAGE	Whether or not fishing alternative is accessed by a popular portage (0,1)
BT*GDLN	Presence of a good boat launch (0,1) times whether trip was taken from boat (1, –1)
BT*NOLN	Presence of no boat launch (0,1) times whether trip was taken from boat (1, –1)
COTTAGE	Presence of significant cottage development (0,1)
LN_UNAC	Natural logarithm of unique access points
W**XXX*	Interaction between attribute *XXX* and whether the angler fished during the winter (1, –1)
MD**XXX*	Interaction between attribute *XXX* and whether the trip was a multiple or day trip (1, –1)

Three measures related to fishing quality were used as attributes in the site choice models. The availability of fish species in a given waterbody was determined if the species was present and was legal to catch and possess at the time of the angling trip. Average reported catch rates for rainbow trout (RT_CUE) were also included for three geographic regions and two times. The use of reported catch rates was a compromise, since there were too few reported catches of rainbow trout (n = 127) to model catch rates by angler, time, and fishing site. The use of an average based estimate for rainbow trout catch rates, however, does result in a biased parameter estimate (Morey and Waldman 1998).

Since almost 80% of the Thunder Bay area anglers prefer to catch walleye (Hunt 2006), we modelled the expected catch of walleye (E(W_CUE)) for each water body and each trip. The parameter estimates from a tobit model of reported walleye catch rates are presented in the Appendix.

We grouped areas for launching boats into no boat launches (NOLN), good boat launches (GDLN), and other boat launches. Good boat launches consisted of concrete, gravel, and sand, while other boat launches included rocks and grass. The boat launch types were interacted with whether the angler did or did not use a boat (BT) on a given fishing trip. Other attributes included the presence of cottage development (COTTAGE) and sites accessible through a portage (PORTAGE).

We accounted for preference heterogeneity in a simple way. Anglers were separated into two groups based on whether or not they reported fishing during the ice season in 2003. Participation in ice fishing trips was thought to provide anglers with information on the presence of fish species and potential summer catch rates at various sites. We investigated whether anglers who ice fished had different preferences for the site attributes by constructing an effects coded variable (W) for participation in winter ice fishing the previous year (+1, –1) and interacted this variable with various attributes. To account for potential differences in preferences for trips of different duration, we used this same approach to examine the effect of day and multiple-day trip contexts. For trip duration, an effects coded variable was created if a trip was a multiple-day trip (MD) (+1, –1) and this was interacted with site attributes. For both types of interactions, we only retained parameter estimates that were statistically significant from zero ($p < 0.05$).

Table 3 presents parameters for the MNL, NL, and CNL models estimated from the 1,152 trips by Thunder Bay area anglers. We estimated all models with GAUSS 7.0 and the MAXLIK subroutine. To limit the chances of identifying a local maxi-

Table 3
Parameter Estimates for the Site Choice Models (Standard Errors in Parentheses)

	MNL		NL		CNL	
Label	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
θ	NA		0.991	(0.054)	0.774 **	(0.040)
OUTSIDE	5.163 **	(0.319)	5.111 **	(0.512)	3.256 **	(0.420)
MD*OUTSIDE	2.655 **	(0.253)	2.650 **	(0.255)	2.446 **	(0.247)
UNKNOWN	5.128 **	(0.308)	5.076 **	(0.506)	3.225 **	(0.411)
MD*UNKNOWN	1.368 **	(0.239)	1.364 **	(0.241)	1.159 **	(0.232)
A_WALL	0.996 **	(0.234)	0.991 **	(0.247)	0.833 **	(0.187)
MD*A_WALL	–0.489 **	(0.187)	–0.485 **	(0.186)	–0.437 **	(0.150)
A_BASS	0.658 **	(0.120)	0.649 **	(0.124)	0.478 **	(0.100)
MD*A_BASS	0.365 **	(0.094)	0.361 **	(0.095)	0.298 **	(0.076)
A_LTROUT	0.877 **	(0.121)	0.866 **	(0.130)	0.699 **	(0.101)
A_BTROUT	1.333 **	(0.176)	1.326 **	(0.195)	0.998 **	(0.147)
A_BSTR	–0.641 **	(0.148)	–0.632 **	(0.149)	–0.536 **	(0.117)
E(W_CUE)	0.881 **	(0.136)	0.875 **	(0.137)	0.649 **	(0.116)
W*E(W_CUE)	–0.173 **	(0.054)	–0.173 **	(0.054)	–0.147 **	(0.045)
MD*E(W_CUE)	0.467 **	(0.128)	0.464 **	(0.128)	0.414 **	(0.103)
RT_CUE	4.310 **	(0.269)	4.283 **	(0.353)	3.304 **	(0.267)
W*RT_CUE	–0.668 **	(0.208)	–0.663 **	(0.208)	–0.538 **	(0.168)
LN_WAREA	0.322 **	(0.032)	0.321 **	(0.037)	0.236 **	(0.029)
T_COST	–0.027 **	(0.002)	–0.027 **	(0.002)	-0.021 **	(0.002)
MD*T_COST	0.016 **	(0.001)	0.016 **	(0.001)	0.013 **	(0.001)
PORTAGE	–1.277 *	(0.507)	–1.270 *	(0.507)	–1.015 *	(0.395)
BT*GDLN	0.729 **	(0.087)	0.724 **	(0.093)	0.579 **	(0.074)
BT*NOLN	–0.809 **	(0.154)	–0.804 **	(0.160)	–0.597 **	(0.125)
COTTAGE	–1.528 **	(0.219)	–1.515 **	(0.227)	–1.103 **	(0.184)
W*COTTAGE	–0.813 **	(0.212)	–0.806 **	(0.214)	–0.637 **	(0.167)
LN_UNAC	0.782 **	(0.095)	0.770 **	(0.104)	0.524 **	(0.084)
LL (β = 0)	–6,988.2		–6,988.2		–6,988.2	
LL (β)	–5,130.5		–5,130.5		–5,120.4	
Adjusted ρ^2	0.261		0.261		0.263	

Notes: All significant tests for dissimilarity parameter estimate (θ) are from one; six alternative specific parameter estimates are not included in the table (results available from authors).
* = significant at the 5% level; ** = significant at the 1% level.

mum for the log likelihood function for the nested and cross-nested logit models, these models were estimated with multiple sets of starting values. The models all provide good fits with the empirical data with adjusted ρ^2 values of about 0.26.

The models are similar in many ways and appeal to intuition. The likelihood of selecting a fishing site increases with decreasing travel costs (T_COST), increasing availability and abundance of desirable fish species (*e.g.*, A_WALL, A_LTROUT, E(W_CUE)), increasing water areas (LN_WAREA), less cottage development (COTTAGE), increasing number of access points (LN_UNACC), and better quality boat launch sites (BT*GDLN). The parameter estimates for the alternative specific constants of UNKNOWN and OUTSIDE account for the 2.2% and 2.5% of modeled fishing trips for these alternatives, respectively.

The significance of the dissimilarity parameter (θ) for the CNL model suggests that spatially close fishing sites share some unobserved utility. Other than the dissimilarity parameter estimate, few differences in the other parameters are apparent among the models. The alternative specific constants for fishing trips to unknown sites (UNKNOWN and MD*UNKNOWN) and to sites outside the study area (OUTSIDE and MD*OUTSIDE) differed, since these alternatives were not part of the spatial nesting structure. Almost all parameter and standard error estimates were smaller for the CNL than for the MNL models.

A likelihood ratio test suggests a significant improvement in model fit for the CNL than the MNL model (χ^2 = 20.21, df = 1, $p < 0.001$). The NL model provided no improvement over the MNL model (χ^2 = 0.02, df = 1, $p = 0.879$). The CNL and NL models are not nested as the allocation parameters (α) were provided rather than estimated. Since these models are based on the same number of attributes and data, the Akaike Information Criterion supports the model with the lower log likelihood. The lower log likelihood for the CNL model (–5,120.4) when compared to the NL model (–5,130.5) provides support for the CNL model.[2]

Management Scenarios

While the CNL model represented a significant statistical improvement over the MNL and NL models, the implications of the CNL model on substitution among fishing sites and welfare measures from management scenarios remain unclear. We developed two management scenarios to assess these implications for the MNL and CNL models. The nested logit model estimates were not included, since the nested logit model did not represent a significant improvement over the MNL model.

The first scenario involved the degradation of a network of logging roads near Thunder Bay area that lead many anglers to fishing sites. The degradation involved changing the speeds on single lane gravel roads with maintenance problems to 7.5 km per hour (*i.e.*, conversion to trails) and all other gravel roads to 30 km per hour (*i.e.*, conversion to a single lane gravel road with maintenance problems). This change in road degradation serves to increase travel costs through the opportunity cost of travel time and would affect the costs of accessing 41 of the 431 fishing alternatives. The second scenario focused on a 50% decline in expected walleye catch rates at fishing sites along a river and a reservoir (affecting five sites).

We first examined differences in patterns of forecasted fishing site choices among the MNL and CNL models for the two scenarios. The forecasts were calculated using parameter estimates from table 3. The predicted probabilities of fishing

[2] Similar findings were observed for analyses based on data from a different population of northern Ontario anglers. These results are available by request from the authors.

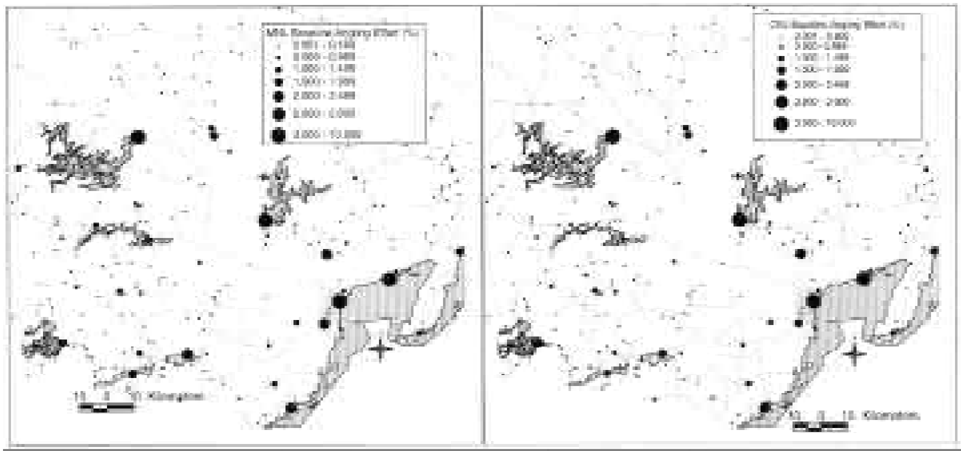

Figure 2. Baseline Predicted Angling Effort to Sites from MNL and CNL Models

site choice before the scenario changes are shown in figure 2 for the MNL and CNL models, respectively. Few differences are observable between the figures, since there are 431 choice alternatives available to the anglers (*i.e.*, expected predicted probabilities of choice among the sites are small). The models predict that fishing site use would be highest among a handful of fishing sites that were either part of Lake Superior or at a number of large, road-accessible lakes containing walleye.

The predicted relative changes in fishing site use from the CNL model are shown in figure 3. For the road degradation scenario (left side of figure 3), the affected sites (the large circles) were predicted to have large relative losses in use. This predicted loss was expected to lead to increased relative use at all other fishing sites. However, the CNL model predicted greater relative increases in use at sites nearer than farther away from the affected sites (*i.e.*, larger triangles at sites near the affected sites). The model predicted that fishing sites near the affected sites were better substitute fishing sites than were sites located further away.

This same spatial pattern was found for the reduced walleye catch rate scenario (right side of figure 3). The CNL model again predicted that the affected sites (large circles) would have large relative losses. The model also predicts that use at all unaffected sites would increase. While the pattern appears weaker than the pattern for the road degradation scenario, fishing sites closer to the affected sites exhibited greater increases in predicted relative use than use at more distant fishing sites.

To compare the forecasts from the MNL and CNL models, a figure was produced from the differences in the relative change in predicted use at sites (*i.e.*, MNL forecasted relative change – CNL forecasted relative change). To avoid the need to consider the sign of the effect, we transformed the predicted relative changes into absolute values.

For the road degradation scenario (left side of figure 4), the MNL model predicts that many affected fishing sites would have larger relative losses in use than those predicted by the CNL model (see the large triangles in the affected area). The CNL model predicts greater relative use at fishing sites in close proximity to the affected sites than would the MNL model (see circles). By contrast, the MNL model predicts increased relative use of fishing sites that were further from the affected sites than did the CNL model (see small triangles). This is the result of the MNL model assuming that changes in attributes at affected sites affect the relative use at all other sites equally.

Similar results were found for the reduced walleye catch rate scenario (right side of figure 4). The MNL model again predicted larger relative losses in use at the affected sites than did the CNL model (see large triangles at affected sites). The CNL model predicted that sites near the affected sites would have greater increases in relative use than those predicted by the MNL model (see circles). Finally, the

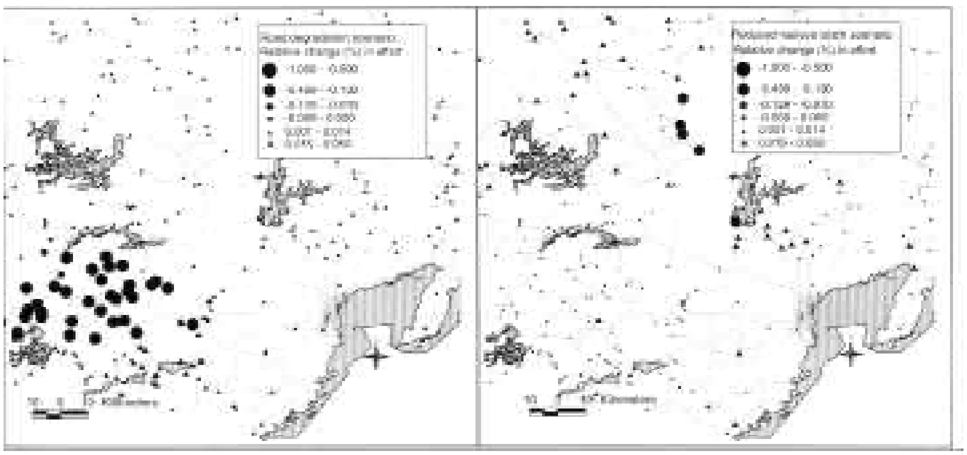

Figure 3. Predicted Relative Change in Effort+ at Fishing Sites from CNL Model for Road Degradation and Declining Walleye Population Scenarios

Note: +– estimated as [Probability $(P)_{\text{before change}} - P_{\text{after change}}] / P_{\text{before change}}$

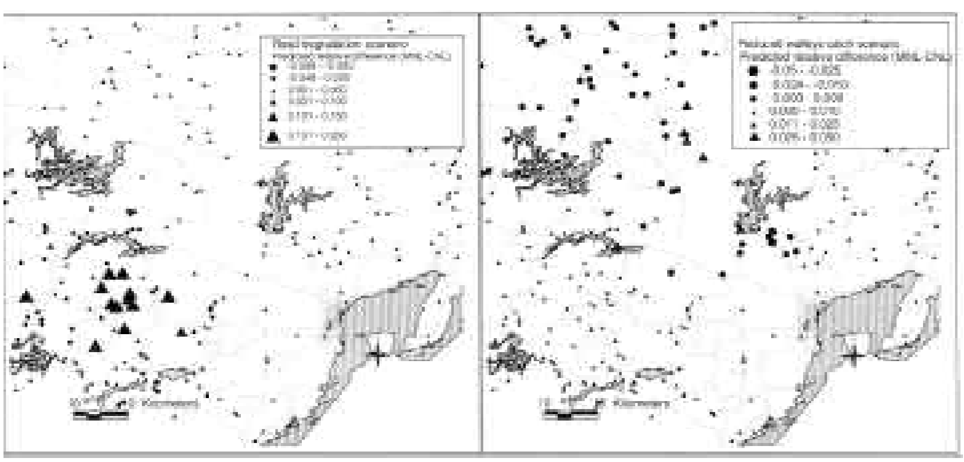

Figure 4. Predicted Difference in Relative Use Change+ from MNL and CNL Models for Road Degradation and Declining Walleye Catch Scenarios

Note: + – estimated as ABS{[Probability $(P)_{\text{MNL before}} - P_{\text{MNL after}}]/P_{\text{MNL before}}\}$ – ABS{$[P_{\text{CNL before}} - P_{\text{CNL after}}]/P_{\text{CNL before}}\}$

MNL model predicted that fishing sites far away from the affected sites would have greater increases in relative use than the CNL model.

A remaining question about the CNL model is the implications of its use to estimate per-fishing-trip CV. We used observed trip data to estimate the median per-trip CV for both day and multiple-day trips by taking 1,000 random draws of the parameter estimates based on the mean values and covariance matrix of these estimates. Following Hagerty and Moeltner (2005), we used this bootstrapping approach to report the mean and 95% confidence interval estimates that represent the expected median per-trip CV.

Both the CNL and MNL models estimate a significant, negative loss in per-trip economic benefits for both day and multiple-day trips for the road degradation scenario (see table 4). We predicted a larger per-trip welfare loss for multiple than day trips (about \$0.33 compared to \$0.24), since the interaction term of multiple-day trips with travel costs (MD*T_COST) was significant and positive. For the CNL model, the welfare measures for both day and multiple-day trips are larger and more variable. However, the mean CV estimate for multiple-day trips from the CNL model falls within the confidence interval of the MNL model, suggesting the two welfare measures are not statistically different. For day trips, the mean CV estimate for the CNL model was greater than the values in the confidence interval for the MNL model.

All CV estimates for multiple-day trips were significant and negative for the reduced walleye catch rate scenario. However, the welfare measures from both models for day trips were not significantly different from zero, in part because the positive walleye catch rate (E(W_CUE) is attenuated for day trips (negative value of M*E(W_CUE)) and for winter anglers (W*E(W_CUE)). The mean per-trip CV estimates for multiple-day trips from the CNL model were again greater and more variable than were the estimates from the MNL model. The CV estimates for day trips from the CNL model, however, were lower and less variable than those estimates from the MNL model. This difference arose partly because of the relative impact that the CNL model had on walleye catch rate estimates. For the MNL

Table 4

Per-trip Estimates of Compensating Variation (CV) for Scenarios Estimated using a Cross-Nested (CNL) and a Multinomial Logit (MNL) Model

	Mean Per-trip Values (95% Confidence Interval)			
	Road Degradation Scenario		Decline to Walleye Catch Rate Scenario	
	MNL	CNL	MNL	CNL
Day trips				
	–\$0.24 (–\$0.26 to –\$0.22)	–\$0.27 (–\$0.30 to –\$0.25)	–\$0.30 (–\$0.59 to \$0.02)	–\$0.17 (–\$0.47 to \$0.06)
Multiple-day trips				
	–\$0.33 (–\$0.36 to –\$0.29)	–\$0.35 (–\$0.39 to –\$0.31)	–\$1.47 (–\$2.93 to –\$0.74)	–\$1.71 (–\$3.67 to –\$0.87)

model, the parameter estimate for the interaction of trip duration and expected walleye catch rate (MD*E(W_CUE)) was 53.0% of the size of the estimate for the expected walleye catch rate (E(W_CUE)). The relative size of this estimate for the CNL model was 63.8%. This difference meant that the effect of expected walleye catch rates for day trips was relatively lower for the CNL than the MNL model (*i.e.*, the parameter estimate for expected walleye catch for day trips equals E(W_CUE) minus MD*E(W_CUE)). These results suggest that the implications on CV from a CNL model are sensitive to the application and the attributes that one investigates.

Conclusions

We employed a CNL model that provides a more flexible treatment of spatial relationships among alternatives than does a NL model. Following Bhat and Guo (2004), we used a CNL model that was tractable to estimate with over 400 alternatives. This number of alternatives is considerably larger than most recreational applications of NL models in the literature, and is a necessary consideration in our empirical application due to the availability of hundreds of fishable waterbodies in the Canadian Shield region of Ontario. The CNL model permits one to overlap the membership of alternatives among multiple nests. By using distance decay functions to allocate an alternative among the various nests, one can test for (and if appropriate) circumvent the IIA property. Unlike multinomial probit and mixed logit models, a CNL model avoids estimation of integrals. While methods exist to estimate models requiring numerical integration (see Train 2003), the use of closed form choice models, such as the CNL model, avoids concerns with simulation error (Bhat and Guo 2004).

In our application, the CNL model outperformed both the MNL and NL models. While the nests for the NL model were based on the same spatial support points as the CNL model, the NL model did not provide a significant improvement over the MNL model. As noted by Herriges and Phaneuf (2002), there is little reason to support the use of a NL when other ways to accommodate correlation among unobserved utilities are available.

The CNL model employed on fishing site choices revealed a much more complex pattern of spatial substitution among fishing sites than either the MNL or NL models. Using two hypothetical management scenarios, the site choice predictions from the CNL model were different than those predictions from the MNL model. The allocation of sites among multiple nests in the CNL model softened the predicted relative impacts from the scenarios on affected sites and strengthened the relative impacts on use at sites near the affected sites.

These forecasting differences among the CNL and MNL models have important implications for practitioners. For example, a fisheries manager using the MNL model may conclude that a closure of a popular fishing site will simply involve absorption of angling trips over all available alternatives. By using the CNL model, the fisheries manager may become concerned about the expected increase in relative use at fishing sites near the closed site.

Researchers use random utility models to assess changes in economic welfare that result from management scenarios. In two management scenarios, we found some differences in the mean and variability of per-trip CV estimates. While in three instances the CV estimates from the CNL model were greater and more variable than those estimates from the MNL model, the fourth case did not follow this pattern. The economic welfare implications of using a CNL model instead of an MNL model appear to defy a simple explanation and are likely to vary among empirical applications.

While this research illustrates the implications of using a basic CNL model to examine recreational fishing site choices, a few issues remain to be resolved. First, our nesting structure used for the CNL model was deterministic. As with the NL model (Herriges and Kling 1997), the choice probability and CV estimates from our scenarios were conditional upon our nesting structure and allocation measures. Second, we restricted the dissimilarity parameters for all nests to be equal. This use of a global dissimilarity parameter may have hidden interesting spatial pockets where neighbouring sites act as better or worse substitutes. Relaxing this restriction would be consistent with the trend in spatial analyses of moving from global towards local models to account for spatial effects (see Fotheringham 1997). Third, we did not assess the implications of using a CNL model on trip participation models that researchers typically estimate along with site choice models. Finally, our treatment of the catch rate for rainbow trout may be problematic (Morey and Waldman 1998). The use of reported catch rates does not account for differences in the abilities of anglers to catch these fish.

References

Ben-Akiva, M.E., and M. Bierlaire. 1999. Discrete Choice Methods and Their Applications to Short-term Travel Decisions. *Handbook of Transportation Science: International Series in Operations Research and Management Science*, R.W. Hall, ed., pp. 5–34. Norwell, MA: Kluwer.

Bhat, C.R., and J. Guo. 2004. A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling. *Transportation Research, Part B* 38:47–68.

Bierlaire, M., K.W. Axhausen, and G. Abay. 2001. The Acceptance of Model Innovation: The Case of the Swissmetro. Monte Verità/Ascona. Presented at the 1st Swiss Transport Research Conference (March 1-3).

Bolduc, D. 1992. Generalized Autoregressive Errors in the Multinomial Probit Model. *Transportation Research, Part B* 26(2):155–70.

___. 1999. A Practical Technique to Estimate Multinomial Probit Models in Transportation. *Transportation Research Part B* 33:63–79.

Bolduc, D., B. Fortin, and M. Fournier. 1996. The Effect of Incentive Policies on the Practice Location of Doctors: A Multinomial Probit Analysis. *Journal of Labor Economics* 14(4):703–32.

Bolduc, D., B. Fortin, and S. Gordon. 1997. Multinomial Probit Estimation of Spatially Interdependent Choices: An Empirical Comparison of Two New Techniques. *International Regional Science Review* 20:77–101.

Canadian Automobile Association. 2004. *Driving Costs:* 2004 *Edition*. World Wide Web publication www.caa.ca/e/automotive/pdf/driving-costs-04.pdf accessed July 15 2005.

Englin, J., P. Boxall, and D. Watson. 1998. Modeling Recreation Demand in a Poisson System of Equations: An Analysis of the Impact of International Exchange Rates. *American Journal of Agricultural Economics* 80:255–63.

Fotheringham, A.S. 1997. Trends in Quantitative Methods I: Stressing the Local. *Progress in Human Geography* 21(1):88–96.

Garrido, R.A., and H.S. Mahmassani. 2000. Forecasting Freight Transportation Demand with the Space-time Multinomial Probit Model. *Transportation Research, Part B* 34:403–18.

Grijalva, T.C., R.P. Berrens, A.K. Bohara, P.M. Jakus, and W.D. Shaw. 2002. Valuing the Loss of Rock Climbing Access in Wilderness Areas: A National-level Random-utility Model. *Land Economics* 78:103–20.

Hagerty, D., and K. Moeltner. 2005. Specification of Driving Costs in Models of Recreation Demand. *Land Economics* 81:127–43.

Hanemann, W.M. 1982. Applied Welfare Analysis with Qualitative Response Models. Working Paper No. 241. University of California, Berkeley, CA.

Hauber, A.B., and G.R. Parsons. 2000. The Effect of Nesting Structure Specification on Welfare Estimation in a Random Utility Model of Recreation Demand: An Application to the Demand for Recreational Fishing. *American Journal of Agricultural Economics* 82:501–14.

Herriges, J.A., and C.L. Kling. 1997. The Performance of Nested Logit Models when Welfare Estimation is the Goal. *American Journal of Agricultural Economics* 79:792–802.

Herriges, J.A., and D.J. Phaneuf. 2002. Inducing Patterns of Correlation and Substitution in Repeated Logit Models of Recreation Demand. *American Journal of Agricultural Economics* 84:1076–90.

Hunt, L.M. 2006. *Modelling Relationships between Road Access and Recreational Fishing Site Choice while Accounting for Spatial Complexities*. Unpublished PhD thesis. Department of Geography and Environmental Studies, Wilfrid Laurier University, Waterloo, ON.

Jones, C.A., and F. Lupi. 1999. The Effect of Modeling Substitute Activities on Recreational Benefit Estimates. *Marine Resource Economics* 14:357–74.

Kling, C.L., and C.J. Thomson. 1996. The Implications of Model Specification for Welfare Estimation in Nested Logit Models. *American Journal of Agricultural Economics* 78:103–14.

McFadden, D. 1978. Modelling the Choice of Residential Location. *Spatial Interaction Theory and Planning Models*, A. Karlqvist, L. Lundqvist, F. Snickars, and J. W. Weibull, eds., pp. 75–96. Amsterdam: North Holland Publishing.

Morey, E.R. 1999. Two RUMs Uncloaked: Nested Logit Models of Site Choice and Nested-logit Models of Participation and Site Choice. *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*, J.A. Herriges, and C.L. Kling, eds., pp. 65–120. Northhampton, MA: Edward Elgar.

Morey, E.R., and D.M. Waldman. 1998. Measurement Error in Recreation Demand Models: The Joint Estimation of Participation, Site Choice, and Site Characteristics. *Journal of Environmental Economics and Management* 35(1):262–76.

Train, K.E. 2003. *Discrete Choice Methods with Simulation*. New York, NY: Cambridge University Press.

Wen, C.-H., and F.S. Koppelman. 2001. The Generalized Nested Logit Model. *Transportation Research, Part B* 35:627–41.

Appendix

Tobit Model Parameter Estimates based on Reported Walleye Catch Rates

Variable	Parameter Estimate	Standard Error	t-value	Probability
Intercept	–3.73086	1.60713	–2.32	0.01
Travel distance from community (km)	0.00281	0.00078	3.63	<0.01
Accessible via 500m or more ATV or walking trail	1.29441	0.22956	5.64	<0.01
Water area (ha)	0.00003	0.00001	4.50	<0.01
Presence of lake trout	–0.20574	0.09157	–2.25	0.01
Presence of smallmouth bass	–0.32951	0.10034	–3.28	<0.01
Primary or secondary targeted species	0.64592	0.18765	3.44	<0.01
Fished from boat	0.74499	0.13947	5.34	<0.01
Motivated to test equipment	0.12964	0.03765	3.44	<0.01
Motivated to relax	0.18666	0.03763	4.96	<0.01
Age	–0.07336	0.03478	–2.11	0.02
Age (square root)	0.92576	0.47156	1.96	0.02
Own a four wheel drive vehicle	0.48343	0.08674	5.57	<0.01
Area (Thunder Bay +1, Wawa –1)	–0.25446	0.05840	–4.36	<0.01
Intercept for Garden Lake	1.30405	0.46896	2.78	<0.01
Intercept for Bedivere Lake	0.79712	0.34862	2.29	0.01
Intercept for Dog River	0.60068	0.23184	2.59	<0.01
Intercept for Poshkokagan and Cheeseman Lakes	0.62590	0.25680	2.44	0.01
Intercept for Nelson, Swallow, and Batwing Lakes	0.71123	0.20453	3.48	<0.01
Intercept for Kagiano Lake	1.16737	0.42830	2.73	<0.01
Sigma	1.44867	0.02742	52.84	<0.01

Source: Hunt (2006).